

DragDiffusion: Harnessing Diffusion Models for Interactive Point-Based Image Editing

Year: 2023 | Citations: 296 | Authors: Yujun Shi, Chuhui Xue, Jiachun Pan, Wenqing Zhang, Vincent Y. F. Tan

Abstract

Accurate and controllable image editing is a challenging task that has attracted significant attention recently. Notably, DRAGGAN developed by Pan et al. (2023) [33] is an interactive point-based image editing framework that achieves impressive editing results with pixel-level precision. However, due to its reliance on generative adversarial networks (GANs), its generality is limited by the capacity of pretrained GAN models. In this work, we extend this editing framework to diffusion models and propose a novel approach Dragdiffusion. By harnessing large-scale pretrained diffusion models, we greatly enhance the applicability of interactive point-based editing on both real and diffusion-generated images. Unlike other diffusion-based editing methods that provide guidance on diffusion latents of multiple time steps, our approach achieves efficient yet accurate spatial control by optimizing the latent of only one time step. This novel design is motivated by our observations that UNet features at a specific time step provides sufficient semantic and geometric information to support the drag-based editing. Moreover, we introduce two additional techniques, namely identity-preserving fine-tuning and reference-latent-control, to further preserve the identity of the original image. Lastly, we present a challenging benchmark dataset called DRAGBENCH■ the first benchmark to evaluate the performance of interactive point-based image editing methods. Experiments across a wide range of challenging cases (e.g., images with multiple objects, diverse object categories, various styles, etc.) demonstrate the versatility and generality of Dragdiffusion. Code and the Dragbench dataset: <https://github.com/Yujun-Shi/DragDiffusion>.